Renewable and Sustainable Energy Reviews 28 (2013) 204–214

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

A time series-based approach for renewable energy modeling

Fatih Onur Hocaoglu [a,b], Fatih Karanfil [c,d,*]

[a] Engineering Faculty, Department of Electrical Engineering, Afyon Kocatepe University, Afyonkarahisar, Turkey
[b] Solar and Wind Energy Research and Application Center, Afyon Kocatepe University, Turkey
[c] EconomiX-CNRS, Department of Economics, University of Paris Ouest, Nanterre, France
[d] Galatasaray University Economic Research Center (GIAM), Istanbul, Turkey

ARTICLE INFO

Article history:
Received 24 March 2013
Received in revised form
15 July 2013
Accepted 20 July 2013
Available online 24 August 2013

Keywords:
Renewable energy
Time series
Prediction

ABSTRACT

Despite the growing literature on renewable energy sources, causal relationships between the variables that are selected as inputs of the models proposed in forecasting studies have not been investigated so far. In this paper, a novel approach to decide prediction input variables of wind and/or temperature forecasting models is suggested. This approach uses time series techniques; more specifically, Granger causality and impulse-response analyses between some meteorological variables. To conduct our study, wind speed, temperature and pressure data obtained from different regions of Turkey are employed. The results suggest that bidirectional causal relationships exist between these variables and that short-run dynamics differ with respect to location (inland versus coastal area). From this, it is concluded that renewable energy models must be built accordingly to improve prediction accuracy.

Contents

1. Introduction

The usage of renewable sources such as solar and wind are exponentially increasing day by day. Accordingly, with the increasing attention being paid by energy industries, governments, non-governmental institutions or the general public to issues relating to the use and development of renewable energy technologies, the research in this field has experienced a remarkable impetus in recent years. As will be discussed later, one of the avenues that has been investigated by researchers is the prediction of renewable energy generation. A major challenge that the scholars in this area have to face is that since meteorological variables carry stochastic behavior, it is not straightforward to determine which variables carry the most important information to predict another one. In other words, whatever the prediction method employed the output of the prediction depends to some extent on some *a priori* assumptions on the relationship between the meteorological variables involved in the analysis. Consider the following example to clarify further this point. From a statistical viewpoint, in a clear region, temperature data and solar radiation data should be strongly correlated. However, in a

* Corresponding author at: University of Paris Ouest, Department of Economics, 200 Avenue de la République, 92001 Nanterre cedex, France.
Tel.: +33 140977815; fax: +33 140974198.
E-mail addresses: hocaoglu@aku.edu.tr (F.O. Hocaoglu), fkaranfil@u-paris10.fr, fatihkaranfil@gmail.com, fatih_9@yahoo.com (F. Karanfil).

1364-0321/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.rser.2013.07.054

cloudy region, the correlation between the two is likely to be less apparent. In this perspective, it becomes clear that, for a specific region, if the existence of a possible relationship between meteorological variables is investigated and in the case of a relationship, if the nature of this relationship can be determined in terms of causality before a prediction model is developed and estimated, then both the accuracy of the model and the performance of the prediction can be improved. That is to say, a pre-analysis method should be developed and used before the prediction step. This is what the present paper aims to achieve. Therefore unlike the previous articles, in this study a pre-analysis methodology is suggested to test for the existence of a relationship between the methodological variables (both predicted variables and those used for prediction) and to explore the causal nature of this relationship.

Essentially the method presented in this study consists of using time series techniques to assess the inter-relationship between a number of variables for renewable energy sources. In fact the use of such techniques has been receiving considerable attention in recent years especially with the increased focus on energy demand forecasting (see [1] for a recent review of this literature). The approach proposed in this paper differs from the exiting literature in that it focuses on both the possible causal chains among the meteorology-related variables and the magnitude of their deviations from the steady-state at each point in time. Our idea is that; a variable (for example, in our case, temperature or pressure) should be used for forecasting another (for example wind speed), only if there is a causality running from the former to the latter. In a seminal paper, Granger argues that this causality technique can be used for determining whether one time series variable is useful in forecasting another and as discussed below, in its general form, it provides a cross-disciplinary framework for studying different relationships in different contexts [2].

It should be noted that the econometric methods used in this study are not new in the economics literature. However, it appears that the idea of using time series techniques with meteorological data is new in renewable energy literature. More specifically, the novelty of this paper is that it shows the way in which the information required for renewable prediction studies can be obtained using Granger causality procedure. The present analysis goes further by providing simulations of generalized impulse response functions (GIRFs) which makes it possible to trace out the transmission mechanisms revealed by the meteorological data used in the analysis. More precisely the GIRFs are used to determine the effect of a shock to any one of the variables on the current and future movements in others. Hence, this final analysis provides further evidence on the short-run dynamic linkages among these variables and how each of them responds to shocks by the other variables in the model.

In view of the above, this paper contributes to two strands of the literature which have been unconnected so far. First, although its main objective is not to make a prediction, it still falls into the literature on renewable prediction since it provides an econometric procedure that enables a coherent model construction for predicting energy resource data and a better understanding of the dynamics of the relationships between the renewable energy sources. In other words, using the causality tests and GIRFs proposed herein, one would be able to decide ex ante on the variables that should be involved in the prediction analysis. Second, the present study adds to the growing body of empirical literature on Granger causality. Using time series analysis as a pre-analysis of renewable prediction, the paper enriches the application fields of Granger causality and GIRFs.

The organization of the rest of the paper is as follows. Section 2 gives a review of previous work on renewable energy prediction. Section 3 is divided into two parts: while in Section 3.1 the data used in the study are presented and discussed, in Section 3.2 the proposed method mentioned above is described and a very brief review of the econometric literature using time series techniques is included. The empirical results are illustrated and analyzed in detail in Section 4. Finally, the conclusion and future works are mentioned in Section 5.

2. Literature review

As mentioned in the previous section, accurate prediction models are of vital importance for energy related issues. Therefore a lot of prediction models developed, tested and published previously. In this section deliberately selected ones are reviewed.

Benghanem et al. developed an artificial neural network (ANN) for daily solar radiation modeling [3]. Soares et al. applied a perceptron neural-network (NN) technique to estimate hourly values of the diffuse solar radiation [4]. Azadeh et al. presented an integrated ANN approach for predicting solar global radiation by meteorological variables [5]. Cao and Lin proposed diagonal recurrent wavelet neural networks to forecast hourly and daily global solar irradiance [6]. In a similar study, Mellit et al. combined wavelet theory and neural networks and proposed a Wavelet-network model to predict daily total solar radiation [7]. In that study, various numbers of total solar radiation data were taken as inputs and different structures were developed. Hocaoglu et al. developed a novel hourly solar radiation forecasting methodology [8]. In a different study, a novel two dimensional modeling procedure is proposed [9]. Since photovoltaic sizing algorithms gives more accurate sizing results in case model generated data used as presented in [10], Hocaoglu developed novel analytical models for photovoltaic based system sizing algorithms and compared modeling performances with previously developed analytical models [11]. In another study, a different ANN based model was developed for the prediction of solar energy potential in Nigeria [12]. In that study meteorological geographical inputs were used in prediction. Behrang et al. used daily mean air temperature, relative humidity, sunshine hours, evaporation, and wind speed values to predict daily global solar radiations [13]. According to an extensive and in-depth review of the literature on the use of artificial intelligence (AI) in photovoltaic system modeling by Mellit et al., AI techniques provide the possibility for sizing PV-systems with reasonable accuracy even in the case of a lack of complete data [14]. A more general review of solar energy modeling and prediction techniques can be found in Khatib et al. [15] and also in Besharat et al. [16].

On the other hand, for wind speed prediction, Salcedo-Sanz et al. proposed a method to improve the accuracy of the wind speed prediction systems based on exploiting diversity in the input data of the neural networks [17]. Kavasseri and Seetharaman used fractional f-Autoregressive Integrated Moving Average (f-ARIMA) models to predict day-ahead wind speed forecast [18]. Louka et al. used Kalman filtering to improve wind speed forecast accuracy [19]. In another study, a history based method is developed to predict short-term ahead wind speed [20]. Mohandes et al. used support vector machines in the same topic [21]. Grassi and Vecchio suggested a two-hidden layer neural network to predict the wind energy output [22]. From the same perspective Haque et al. proposed soft computing models (SCMs) “augmented” by a similar days (SD) method which is found to lead to an increase in the level of performance of wind speed forecasting technique [23] (see also [24] for a review of the use of ANN in this field). In a different study, temperature values are employed as model inputs and accurate wind prediction results are obtained using hidden Markov models [25]. Cadenas and Rivera compared ARIMA and ANN models in wind speed predictions [26].

In most of the previous studies, the prediction model inputs are decided intuitively. In this work, the necessity of a pre-analysis for prediction models is emphasized and a novel approach is suggested. This approach consists of using a time series analysis on

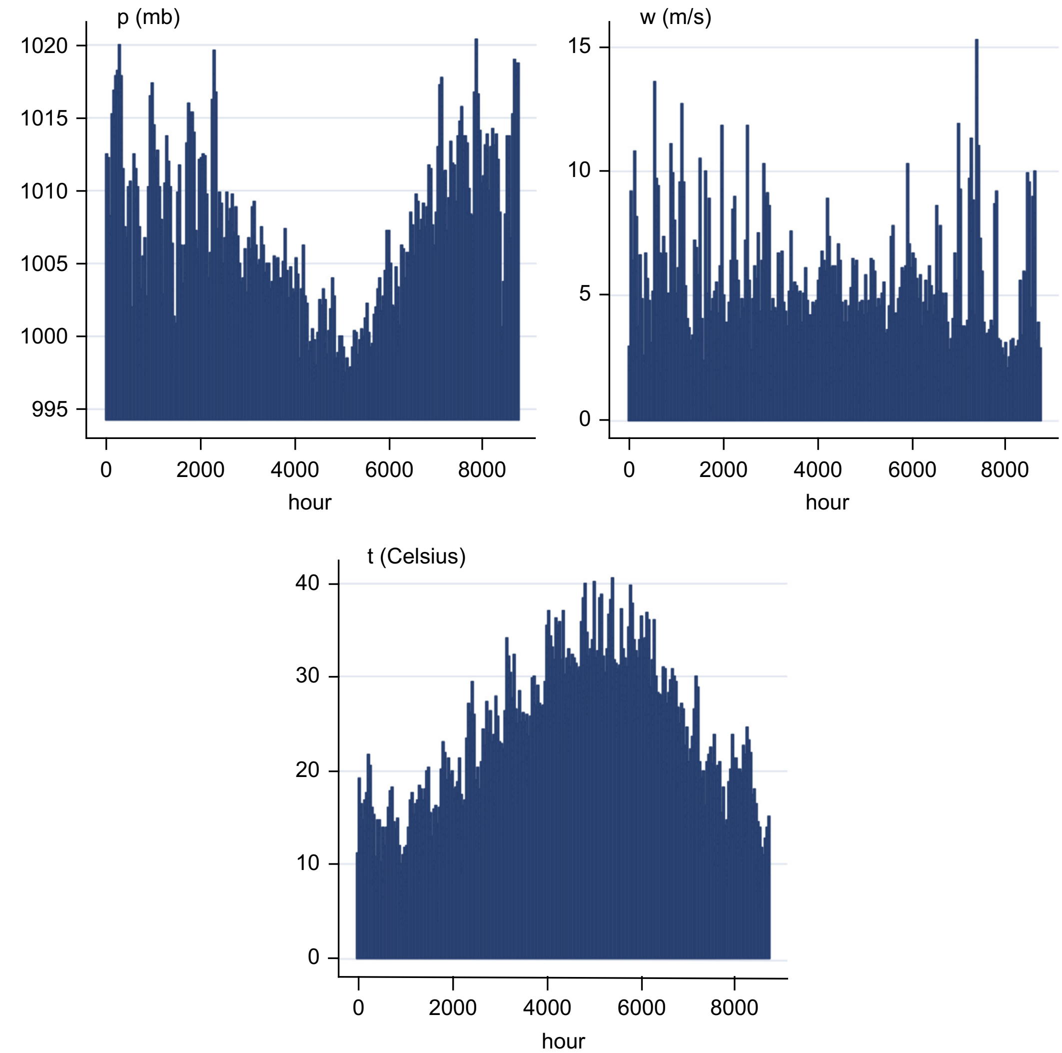

Fig. 1. Temperature, wind speed and air pressure data measured at Antalya region.

some meteorological variables in order to investigate time series properties of each of these variables and then to search for a causal relationship among them. This approach can provide clear indication of which variables should be incorporated in any prediction model as an input, and in consequence, more accurate prediction models can be developed.

3. Data and methodological issues

3.1. Data description

To develop the econometric model, which will be discussed in the following section, hourly measured yearly wind speed, temperature and air pressure data obtained from Antalya, Izmir and Kayseri regions are selected. To prevent the developed approach from possible errors, the measured data are first pre-analyzed. This pre-analysis includes the followings: the missing data that are absent in the whole set are interpolated using future and previous data. Hence, erroneous data with invalid values were rectified as a pre-analysis.

The wind speed, temperature and air pressure data are illustrated in Figs. 1–3 for Antalya, Izmir and Kayseri regions, respectively. The range of data is shown in Table 1.

It is obvious from Figs. 1–3 and Table 1 that, not only the ranges but also the variations of the data are different. Hence, the deliberately selected data has different characterizations.

3.2. Methodology

In recent years, the use of Granger causality procedure have gained increased popularity and today it is performed frequently in a widespread field of applications such as economics or epilepsy analysis.[1] In some promising studies, Granger causality tests are employed to increase our understanding of climate variability. To give some examples of the previous works close to our own, Mosedale et al. provide an application of Granger causality in an ocean–atmosphere system and study the influence of daily sea surface temperatures (SSTs) on daily values of the North Atlantic Oscillation (NAO) [30]. They conclude that there exists a significant feedback of SSTs on the NAO. In another interesting and related paper, Smirnov and Mokhov investigate the effects of variations in carbon dioxide atmospheric content, solar activity, and volcanic

[1] Here are some examples for the use of the Granger method in different areas of research. The most related literature to the present study using Granger causality tests investigates causal relationships between environmental variables (e.g. CO_2 emissions) or energy variables (e.g. primary energy consumption) and economic development indicators (such as gross domestic product (GDP) or industrial value added). See Ref. [27] for an overview of the literature in this field. On the other hand, in the more distant literature, in epilepsy studies for instance, Granger causality tests are generally used for determining direction of relationships between local field potentials (see for an example [28]) or for examining structures of different neurological states (see for example Ref. [29]).

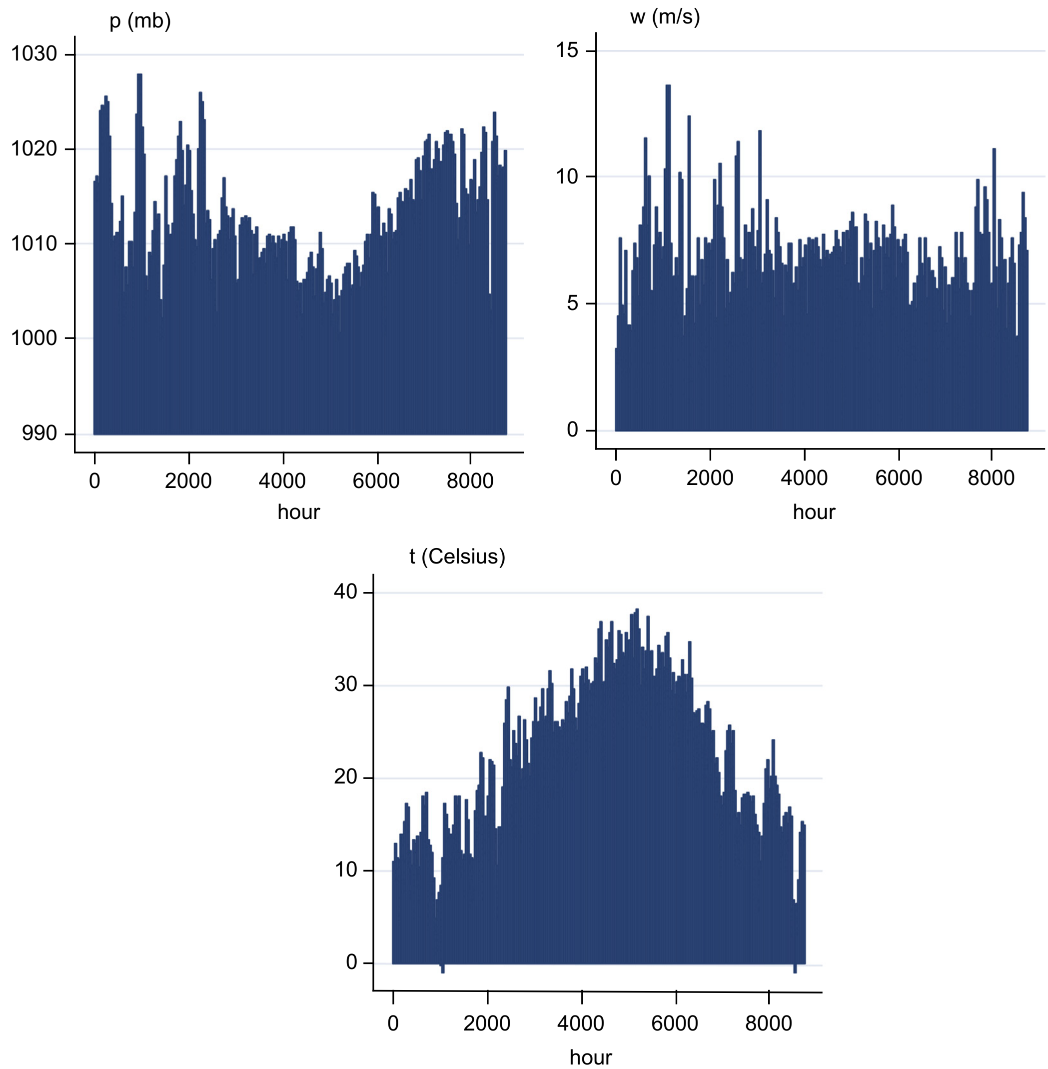

Fig. 2. Temperature, wind speed and air pressure data measured at Izmir region.

activity on the global surface temperature (GST) [31]. They show that all the three factors *Granger cause* the GST.

Since we have given a description of the variables involved and the basic structure of our idea, we now proceed with our econometric analysis which consists of testing for causal relationships and dynamics between the variables of interest, namely air pressure (p), temperature (t) and wind speed (w). It is well known that the order of integration of all variables used in the analysis has to be established before deciding on which time series method to use. For example, if the series are found to be stationary, the standard Granger causality test in an unrestricted vector autoregressive (VAR) model can be conducted by testing whether parameters of explanatory variables are jointly zero. On the other hand, in the same framework, non-stationarity series having a unit root is a random walk and cannot be used for forecast. This problem has given rise to alternative testing procedures such as error-correction model (ECM) [32,33] and autoregressive distributed lags (ARDL) bounds test [34].[2]

In line with the existing literature on time series analysis, first of all, time series properties of the data are checked by performing the augmented Dickey–Fuller (ADF; [37]) and the Phillips and Perron (PP; [38]) unit root tests based on the model given by the following equation:

$$\Delta X_t = \alpha_0 + \beta\mu + \alpha_1 X_{t-1} + \sum_{i=1}^{k} \lambda_i \Delta X_{t-i} + u_t \tag{1}$$

where X is the variable to be tested, μ is the trend variable, Δ is the first-difference operator and u_t is Gaussian white noise. On the other hand, β is the coefficient on the time trend and k is the lag order of the autoregressive process. For any given time series, if $\alpha_1 = 0$ then this series can be characterized as unit root or random walk. In other words, in this case the variable behaves as a non-stationary process. As mentioned above, in order to address the non-stationarity issue, several methods have been suggested in the literature. However, if all variables are found to be stationary, this result will bring us the standard Granger causality framework, since it is not necessary to use more complicated models

Granger gives the definition of causality in terms of prediction: given to time series x and y, "if y_t causes x_t, then x_{t+1} is better forecast if the information in y_{t-j} is used than if it is not used" [39, p. 200]. In time series analysis, this means that one may determine whether one time series (here the variable y) is useful in forecasting another (here the variable x). In our case, this statement implies that the Granger causality procedure can be

[2] Detailed presentation of the time series analysis and Granger causality procedure can be found in Hamilton [35, Chapters 11 and 19]. For a brief discussion on the same issue, see for instance [36].

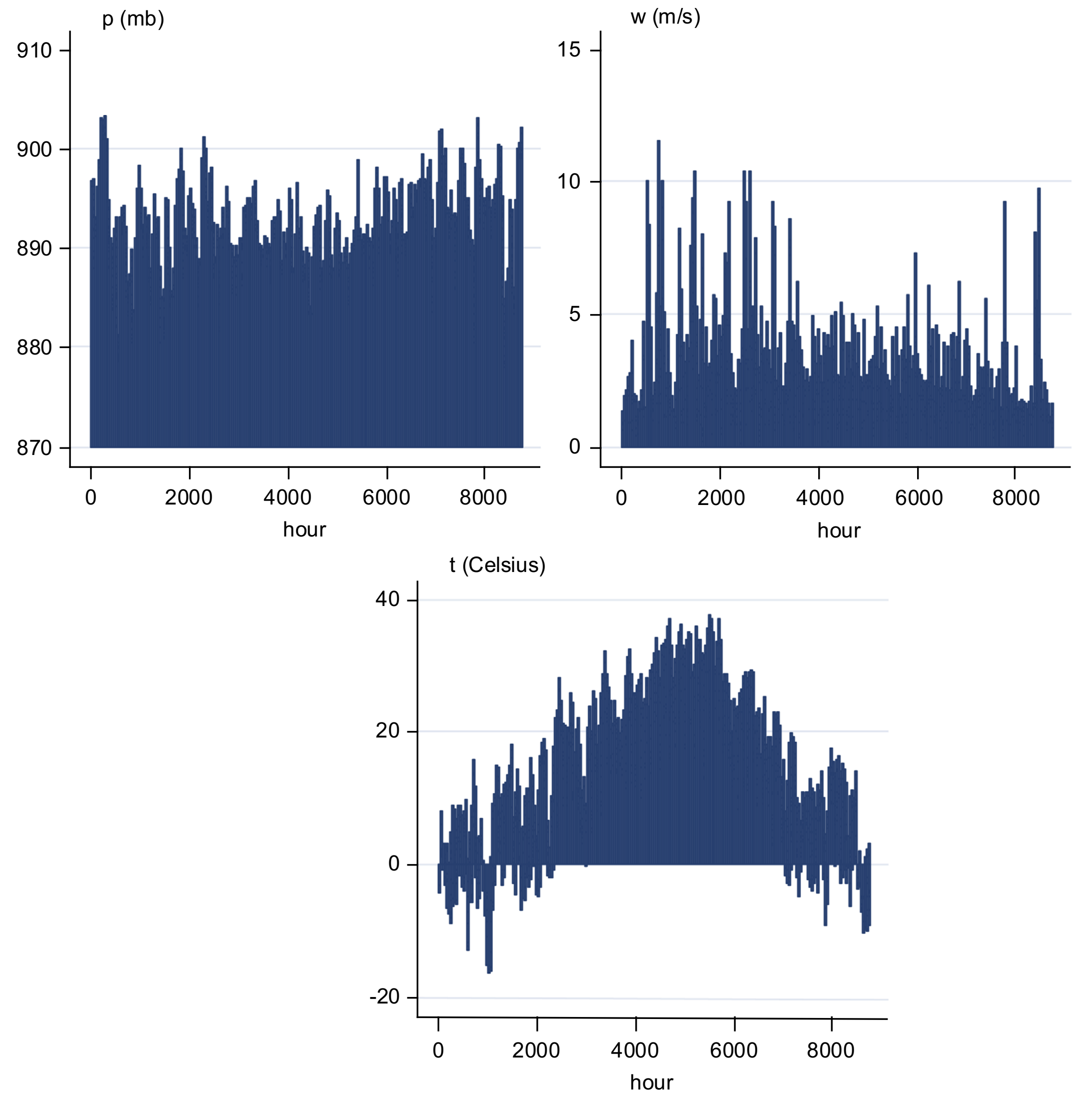

Fig. 3. Temperature, wind speed and air pressure data measured at Kayseri region.

Table 1
Range of the data.

Regions	Temperature (°C)	Wind speed (m/s)	Air pressure (mb)
Antalya	1.3–40.06	0–15.3	994–1020
Izmir	−0.8–38.2	0–13.6	995–1028
Kayseri	−16.4–37.5	0–11.5	873–903

Table 2
Granger causality test hypotheses.

Direction of causality	Test thypotheses
X_1 does not cause X_2	$H_0: \ b_{1i}=0 \quad \forall i=1,\dots,m$ $H_1: \ \exists b_{1i}\neq 0 \quad \forall i=1,\dots,m$
X_1 does not cause X_2	$H_0: \ a_{2i}=0 \quad \forall i=1,\dots,n$ $H_1: \ \exists a_{2i}\neq 0 \quad \forall i=1,\dots,n$

Table 3
VAR Granger causality test results (Block exogeneity Wald test).

Independent variable: air pressure (p)		Independent variable: wind speed (w)		Independent variable: temperature (t)	
Direction of causality	Chi-square	Direction of causality	Chi-square	Direction of causality	Chi-square
Estimation results for Antalya region (included lag number: 4)					
$w \to p$	65.3*	$p \to w$	60.2*	$p \to t$	230.6*
$t \to p$	366.9*	$t \to w$	175.9*	$w \to t$	99.5*
Estimation results for Izmir region (included lag number: 7)					
$w \to p$	58.9*	$p \to w$	89.3*	$p \to t$	256.5*
$t \to p$	518.7*	$t \to w$	479.7*	$w \to t$	192.3*
Estimation results for Kayseri region (included lag number: 6)					
$w \to p$	38.2*	$p \to w$	122.8*	$p \to t$	100.8*
$t \to p$	564.6*	$t \to w$	297.3*	$w \to t$	123.6[a]

* Significance at 1% level.

well suited for wind speed, temperature and pressure forecasting. Before we go on to discuss the econometric methodology, to avoid any possible confusion due to the use of the term "prediction" herein, we would like to underline the following point, although it has been already discussed above and will be emphasized again in the following sections: Granger's method for detecting causality between two or more variables cannot be used to predict any variable. Instead, indicating the direction of causality (if some causal relationships exist), it makes it possible to ascertain which

variable may be used to predict the other. In consequence, the procedure proposed in this paper should be considered as a pre-analysis method for renewable prediction. The variables among which causal relationships are detected may be used ex post to predict energy resource data.

Technically speaking one can show that the process X_1 does not Granger cause the process X_2 if

$$E(X_{2t}/I_{t-1}(X_2), I_{t-1}(X_1)) = E(X_{2t}/I_{t-1}(X_2)) \quad (2)$$

where $I_{t-1}(X_i)$ is the space generated by the linear combinations of the past values of X_i.

With stationary series, following Granger, this can be done by performing a set of tests based on the regressions given below [2]:

$$\begin{aligned} X_{1t} &= \alpha + \sum_{i=1}^{m} a_{1i}X_{1t-i} + \sum_{i=1}^{n} a_{2i}X_{2t-i} + \varepsilon_{1t} \\ X_{2t} &= \beta + \sum_{i=1}^{m} b_{1i}X_{1t-i} + \sum_{i=1}^{n} b_{2i}X_{2t-i} + \varepsilon_{2t} \end{aligned} \quad (3)$$

where α and β are constant terms, ε_{1t} and ε_{2t} are white-noise series and m and n represent the lag orders. The classical Fisher test can be used to detect the causal chains between X_1 and X_2. Granger causality test is then formulated as the hypothesis problems given in Table 2.

We should also mention that this bivariate framework can be extended including more variables in the regression equations. Moreover, it is better to do so, since, as Lutkepohl argued in [40], Granger causality test in a bivariate framework may be subject to the omitted variables bias. Multivariate systems become widely used in some recent studies on time series econometrics. In our model we use temperature, wind speed and, air pressure in order to detect all possible causal chains between them. The results are given in the following section.

4. Empirical results

As we have discussed above, the first step of our empirical analysis consists of testing for the stationarity of the variables involved in our study. Using Eq. (1), ADF and PP unit root tests suggest that all series are stationary, in other words, they are all integrated of order 0, that is $I(0)$ (see Appendix A for the results of unit root tests). In this case, the standard Granger causality test should be employed in a multivariate VAR framework and the system of equations given in Eq. (3) evolves to the following system:

$$p_t = \psi_1 + \sum_{i=1}^{a} \beta_{11i}p_{t-i} + \sum_{i=1}^{b} \beta_{12i}w_{t-i} + \sum_{i=1}^{c} \beta_{13i}t_{t-i} + u_{1t} \quad (4.1)$$

$$w_t = \psi_2 + \sum_{i=1}^{a} \beta_{21i}p_{t-i} + \sum_{i=1}^{b} \beta_{22i}w_{t-i} + \sum_{i=1}^{c} \beta_{23i}t_{t-i} + u_{2t} \quad (4.2)$$

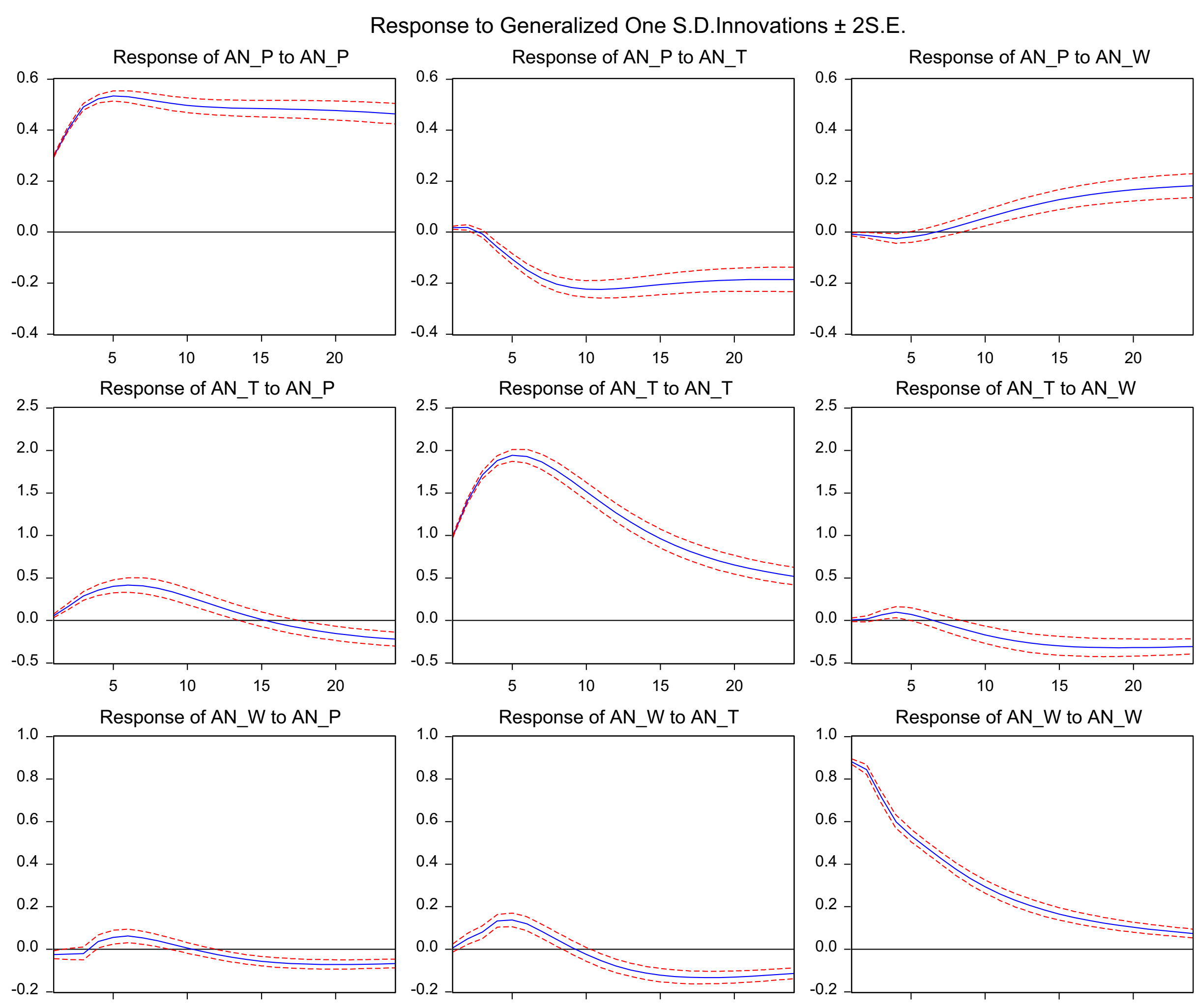

Fig. 4. Generalized impulse-response functions for Antalya.

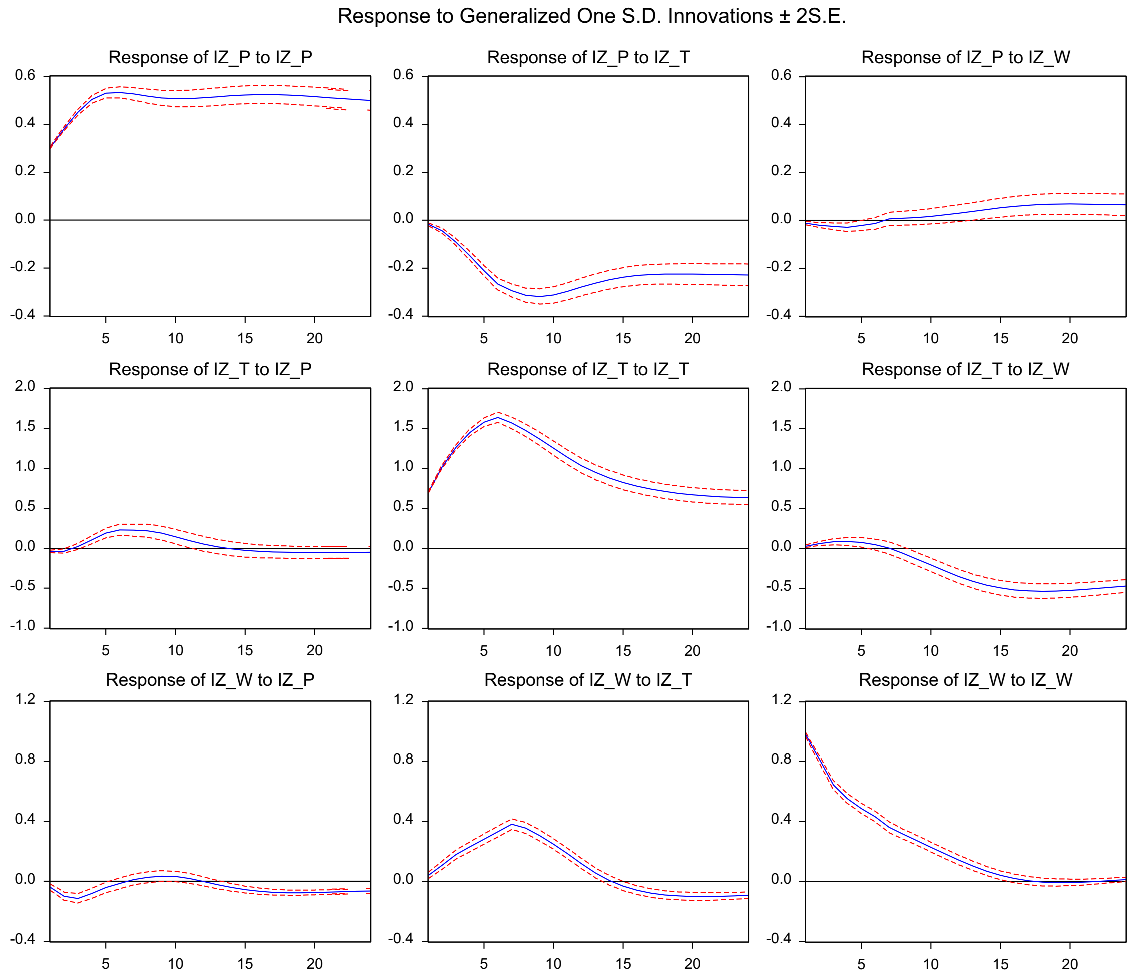

Fig. 5. Generalized impulse-response functions for Izmir.

$$t_t = \psi_3 + \sum_{i=1}^{a} \beta_{31i} p_{t-i} + \sum_{i=1}^{b} \beta_{32i} w_{t-i} + \sum_{i=1}^{c} \beta_{33i} t_{t-i} + u_{3t} \tag{4.3}$$

where all notations (depending on where they are used) are similar to those in Eq. (3). Lag orders a, b and c are selected based on the lag exclusion tests. Significance of the parameters β_i for each variable in each equation will indicate the direction of causality.

We can now apply the standard Wald test for this purpose. The results are reported in Table 3.

From Table 3, it can be seen that: (1) all parameters are found to be significant at the 1% significance level; (2) the neutrality hypothesis (no causal relationship exists) can be rejected at the 1% level; (3) this result holds for all three regions studied and all variables involved without any exception. More specifically, the significance of all parameters in the system of equations (given by Eqs. (4.1)–(4.3)) implies that there exists bidirectional causality among these variables.[3] Taking all together, these findings indicate that if one attempts to conduct a renewable prediction study, temperature, air pressure and wind speed can all be used to do so.

The final step of our empirical work consists of plotting GIRFs of each of the variables for each of the cities. In fact, this method, first introduced by Koop et al. [41] and Pesaran and Shin [42], enables one to assess how each variable responds to innovations in other variables, and to determine whether the shock is permanent or not.[4] Figs. 4–6 show the responses of pressure, temperature, and wind speed to one standard deviation shocks to other variables in the VAR system, including the plus and minus two standard deviations from the mean (representing thus the 95% confidence level).

Figs. 4–6 present the impulse-response paths of the variables involved in the VAR model up to 24 h after a one standard deviation shock is stimulated from others. Overall, comparing the GIRFs for each of the three cities we find very similar results with some differences in sensitivity. It follows also from these figures that there is a considerable self-response (i.e. response to own shock) for all the three variables in the model. Only for the case of wind speed, although it increases in response to its own shock, it moves smoothly downward to its pre-shock level. The oscillating trajectories between different variables when temperature is considered come from the fact that the GIRFs are computed over the entire 24-hour period. However, note also that the level and the period of these oscillations are different with respect to impulse-response variables. For instance, in Antalya (Kayseri) temperature tends to have positive (negative) responses to the shock of pressure at the initial period and then oscillates about zero (dies out). On the other hand, during the 24-hour period, the response of wind speed to a positive one standard deviation shock

[3] We have conducted the same analysis including solar radiation instead of temperature and found identical results to those reported here (see Appendix B).

[4] We prefer to use GIRFs since this method has the advantage over traditional impulse response analyses in that it is invariant to the ordering of the variables in the VAR model.

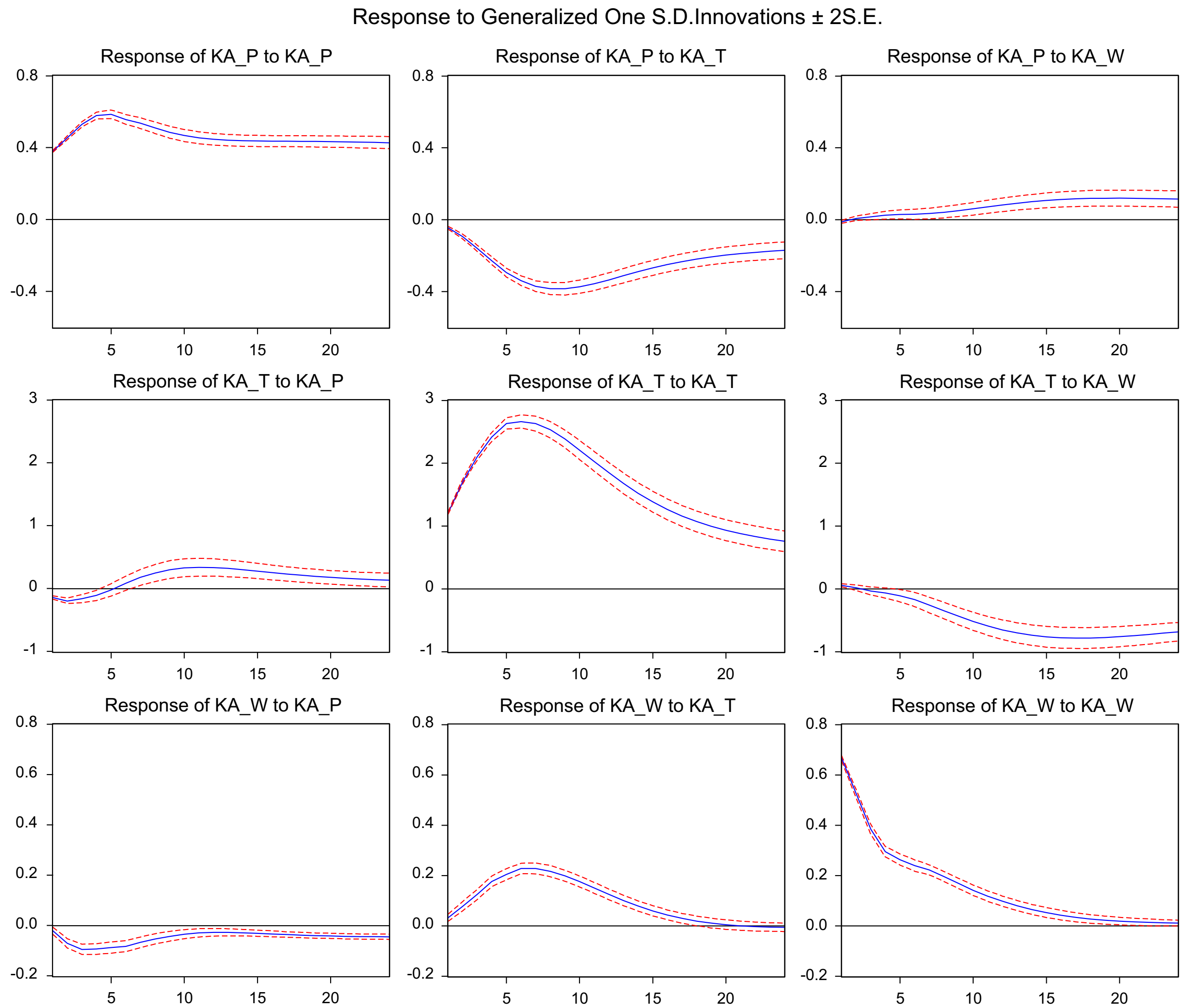

Fig. 6. Generalized impulse-response functions for Kayseri.

Table A1
Results of unit root tests for air pressure, temperature and wind speed.

Variables	Augmented Dickey–Fuller (ADF) test			Phillips–Perron (PP) test		
	Antalya	Izmir	Kayseri	Antalya	Izmir	Kayseri
p	−5.47	−5.40	−7.53	−5.04	−5.66	−6.81
t	−2.39	−3.27	−3.02	−6.39	−3.96	−6.43
w	−11.48	−12.14	−12.83	−22.76	−25.53	−26.48

Critical values for the 1%, 5% and 10% levels of significance are −3.43, −2.86 and −2.56, respectively. Lag determination for the ADF test is based on the Akaike Information Criterion (AIC).

Table B1
Results of unit root tests for radiation.

Variable	Augmented Dickey–Fuller (ADF) test			Phillips–Perron (PP) test		
	Antalya	Izmir	Kayseri	Antalya	Izmir	Kayseri
r	−5.45	−5.02	−5.76	−13.79	−13.72	−14.78

Critical values for the 1%, 5% and 10% levels of significance are −3.43, −2.86 and −2.56, respectively. Lag determination for the ADF test is based on the Akaike Information Criterion (AIC).

Table B2
Causality test results for the VAR system involving radiation instead of temperature.

Independent variable: air pressure (p)		Independent variable: wind speed (w)		Independent variable: radiation (r)	
Direction of causality	Chi-square	Direction of causality	Chi-square	Direction of causality	Chi-square
Estimation results for Antalya region (included lag number: 7)					
$w \to p$	107.1*	$p \to w$	32.4*	$p \to r$	937.9*
$r \to p$	654.4*	$r \to w$	589.6*	$w \to r$	62.4*
Estimation results for Izmir region (included lag number: 5)					
$w \to p$	23.1*	$p \to w$	37.2*	$p \to r$	435.1*
$r \to p$	827.4*	$r \to w$	534.4*	$w \to r$	72.4*
Estimation results for Kayseri region (included lag number: 6)					
$w \to p$	62.3*	$p \to w$	174.5*	$p \to r$	338.1*
$r \to p$	985.8*	$r \to w$	242.2*	$w \to r$	63.0*

* Significance at 1% level.

in the pressure level never becomes positive for the case of Kayseri, contrary to the cases of Antalya and Izmir, which are coastal regions. Another difference that can be observed between coastal and inland regions is that wind speed tends to have a positive impact on the temperature in Antalya and Izmir while the

most evident negative impact of wind speed on temperature is found to be in Kayseri, which is an inland region. These findings indicate that regional differences may lead to different short-run dynamics and impulse-response characteristics. Finally it should be added that the estimated responses to all shocks are transitory and take 40–240 h to die out (see Fig. C1 in Appendix C) and that of the nine shocks for each of the cities, the self-response of pressure is estimated to be the most persistent.

The previous literature on renewable energy prediction has not examined the causal chains between the variables before conducting a prediction study. Among the studies examined in this paper, for instance in [5] seven different variables were considered as inputs of the model to predict solar global radiations. In [13] on the other hand, six different combinations of the input variables were employed to predict global solar radiations. However in these and similar studies, prior to model construction no causality analysis has been undertaken. It is evident that such a task proposed in this study will improve the accuracy of the forecast. Furthermore it may be possible to assign some weights to the variables of the model according to their impulse-response functions in order to develop advanced prediction models. Such analyses should be considered as future tasks of research in this field.

5. Conclusions and future work

In this study, a novel methodology is developed for renewable energy analysis. This methodology suggests a pre-analysis using Granger causality and GIRFs before construction of a prediction model. Using this methodology, the inputs of the prediction model should be determined and unnecessary inputs should be removed in order not to affect negatively the accuracy of the prediction model. In other words, the prediction performance and accuracy can be improved by taking into account the direction of causality among the variables that are to be introduced into the prediction

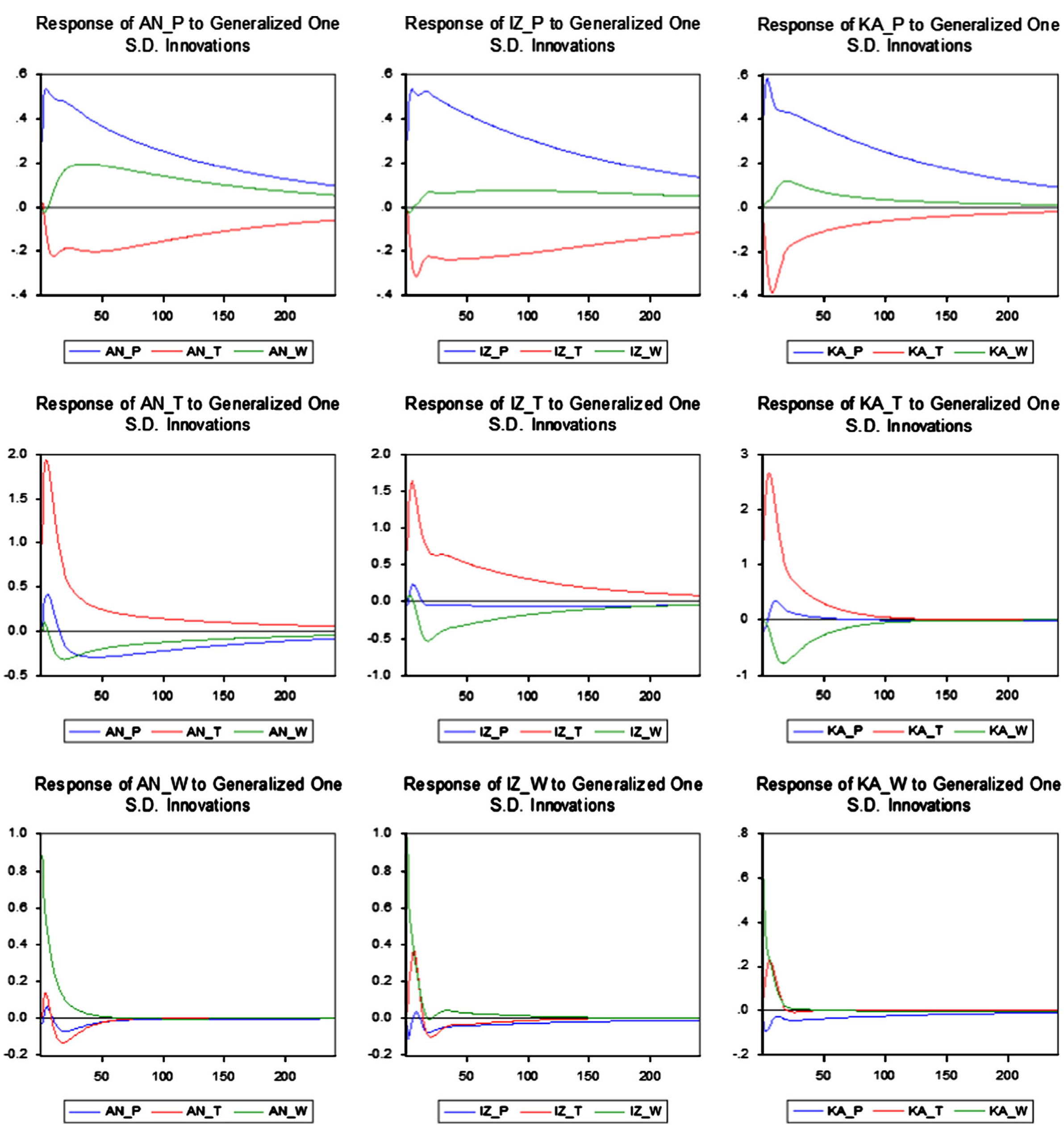

Fig. C1. Generalized impulse-responses up to 240 h.

model. According to our results, bidirectional causal relationships can be established. Hence, taking into account differences in the impulse-response functions, we conclude that although it is possible to use all meteorological variables, in at least these three cities, regional differences (coastal versus inland region) may influence short-run dynamics between the meteorological variables which should be, thus, considered in any prediction study. This confirms our initial suggestion that the accurate variables that should be incorporated in the model must be selected using the suggested methodology.

Future research should take into account regional disparities or differences in periods and intervals of time. Recall that this paper uses hourly data for one year, whereas it is possible that seasonal effects play a strong role in the relationship between the meteorological variables. Accordingly, in different seasons, different causal linkages and short-run dynamics may be found, implying that in prediction studies, different sets of variables may be used in different time intervals. We have not addressed the issue of seasonality in this paper which may prove fruitful as a future research topic.

Acknowledgments

The authors acknowledge the Turkish State Meteorological Service (DMI) for the supply of hourly data.

Appendix A. Unit root tests

Table A1 presents the ADF and PP test statistics for the levels of the variables with the null hypothesis of non-stationarity.

The ADF test results indicate that all the series involved in the analysis are stationary, with the exception of the temperature variable t in Antalya. However, the PP test shows that all variables, including t in Antalya, are stationary even at the 1% level. In fact, it is commonly suggested that using the PP unit root test has several advantages over the ADF test since it does not require a lag length selection procedure and that the error term in the test equation is corrected for both serial correlation and heteroscedasticity [38].

Appendix B. Estimates of the VAR model including radiation series instead of temperature series

The purpose of this appendix is to show that when we substitute temperature series (t) by solar radiation series (r) in the initially estimated multivariate VAR model (see Eqs. (4.1)–(4.3)), the causality tests do not lead to different conclusions. For this, following the procedure already explained in the methodology section, first of all, unit root tests should be conducted to determine the order of integration of the variable r (Table B1).

As was the result for the previous unit root tests for the variables w, p and t, both the ADF and PP tests indicate that the variable r is also a stationary series. Then the causality tests can be carried out by a VAR model composed of three variables (air pressure (p), wind speed (w) and solar radiation (r) instead of temperature (t)). The results are reported in Table B2.

Form Table B2 it follows that bidirectional causal linkages exist between these three variables, as it is the case for the air pressure–wind speed–temperature nexus.

Appendix C. The GIRFs of the regions up to 240 h

See Fig. C1.

References

[1] Suganthi L, Samuel AA. Energy models for demand forecasting – a review. Renewable and Sustainable Energy Reviews 2012;16:1223–40.
[2] Granger CWJ. Investigation causal relations by econometric models and cross-spectral methods. Econometrica 1969;37:424–38.
[3] Benghanem M, Mellit A, Alamri SN. ANN-based modelling and estimation of daily global solar radiation data: a case study. Energy Conversion and Management 2009;50(7):1644–55.
[4] Soares J, Oliveira AP, Bonar MZ, Mlakar P, Escobedo JF, Machado AJ. Modeling hourly diffuse solar-radiation in the city of São Paulo using a neural-network technique. Applied Energy 2004;79:201–14.
[5] Azadeh A, Maghsoudi A, Sohrabkhani S. An integrated artificial neural networks approach for predicting global radiation. Energy Conversion and Management 2009;50(6):1497–505.
[6] Cao J, Lin X. Study of hourly and daily solar irradiation forecast using diagonal recurrent wavelet neural networks. Energy Conversion and Management 2008;49(6):1396–406.
[7] Mellit A, Benghanem M, Kalogirou SA. An adaptive wavelet-network model for forecasting daily total solar-radiation. Applied Energy 2006;83:705–22.
[8] Hocaoglu FO, Gerek ON, Kurban M. Hourly solar radiation forecasting using optimal coefficient 2-D linear filters and feed-forward neural networks. Solar Energy 2008;82(8):714–26.
[9] Hocaoglu FO, Gerek ON, Kurban M. Novel models for hourly solar radiation using a 2-D approach. International Journal of Innovative Computing, Information and Control 2009;5 (1-ICONIP07-03).
[10] Hocaoglu FO, Gerek ON, Kurban M. The effect of model generated solar radiation data usage in hybrid (wind–PV) sizing studies Energy. Conversion and Management 2009;50(12):2956–63.
[11] Hocaoglu FO. Novel analytical hourly solar radiation models for photovoltaic based system sizing algorithms. Energy Conversion and Management 2010;51:2921–9.
[12] Fadare DA. Modelling of solar energy potential in Nigeria using an artificial neural network model. Applied Energy 2009;86:1410–22.
[13] Behrang MA, Assareh E, Ghanbarzadeh A, Noghrehabadi AR. The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. Solar Energy 2010;84:1468–80.
[14] Mellit A, Kalogirou SA, Hontoria L, Shaari S. Artificial intelligence techniques for sizing photovoltaic systems: a review. Renewable and Sustainable Energy Reviews 2009;13:406–19.
[15] Khatib T, Mohamed A, Sopian K. A review of solar energy modeling techniques. Renewable and Sustainable Energy Reviews 2012;16:2864–9.
[16] Besharat F, Dehghan AA, Faghih AR. Empirical models for estimating global solar radiation: a review and case study. Renewable and Sustainable Energy Reviews 2013;21:798–821.
[17] Salcedo-Sanz S, Pérez-Bellido ÁM, Ortiz-García EG, Portilla-Figueras A, Prieto L, Correoso F. Accurate short-term wind speed prediction by exploiting diversity in input data using banks of artificial neural networks. Neurocomputing 2009;72:1336–41.
[18] Kavasseri RG, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. Renewable Energy 2009;34:1388–93.
[19] Louka P, Galanis G, Siebert N, Kariniotakis G, Katsafados P, Pytharoulis I, et al. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. Journal of Wind Engineering and Industrial Aerodynamics 2008;96:2348–62.
[20] Hocaoglu FO, Fidan M, Gerek ON. Mycielski approach for wind speed prediction. Energy Conversion and Management 2009;50:1436–43.
[21] Mohandes MA, Halawani TO, Rehman S, Hussain AA. Support vector machines for wind speed prediction. Renewable Energy 2004;29:939–47.
[22] Grassi G, Vecchio P. Wind energy prediction using a two-hidden layer neural network. Communications in Nonlinear Science and Numerical Simulation 2010;15:262–2266.
[23] Haque AA, Mandal P, Kaye ME, Meng J, Chang L, Senjyu T. A new strategy for predicting short-term wind speed using soft computing models. Renewable and Sustainable Energy Reviews 2012;16:4563–73.
[24] Kalogirou SA. Artificial neural networks in renewable energy systems applications: a review. Renewable and Sustainable Energy Reviews 2001;5:373–401.
[25] Hocaoglu FO, Gerek ON, Kurban M. A novel wind speed modeling approach using atmospheric pressure observations and hidden Markov models. Journal of Wind Engineering and Industrial Aerodynamics 2010;98:472–81.
[26] Cadenas E, Rivera W. Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA–ANN model. Renewable Energy 2010;35:2732–8.
[27] Karanfil F. How many times again will we examine the energy-income nexus using a limited range of traditional econometric tools? Energy Policy 2009;37:1191–4.
[28] Cadotte AJ, DeMarse TB, Mareci TH, Parekh MB, Talathi SS, Hwang D-U, et al. Granger causality relationships between local field potentials in an animal model of temporal lobe epilepsy. Journal of Neuroscience Methods 2010;189:121–9.
[29] Korurek M, Ozkaya A. A new method to estimate short-run and long-run interaction mechanisms in interictal state. Digital Signal Processing 2010;20:347–58.
[30] Mosedale TJ, Stephenson DB, Collins M, Mills TC. Granger causality of coupled climate processes: ocean feedback on the north Atlantic oscillation. Journal of Climate 2006;19:1182–94.

[31] Smirnov D, Mokhov I. From Granger causality to long-term causality: application to climatic data. Physical Review E Statistical, Nonlinear, and Soft Matter Physics 2009;80:016208.
[32] Engle RF, Granger CWJ. Cointegration and error correction: representation, estimation, and testing. Econometrica 1987;55:251–76.
[33] Toda HY, Phillips PCB. Vector autoregressions and causality: a theoretical overview and simulation study. Econometric Review 1994;13:259–85.
[34] Pesaran MH, Shin YC, Smith R. Bound testing approaches to the analysis of level relationships. Journal of Applied Econometrics 2001;16:289–326.
[35] Hamilton JD. Time series analysis. New Jersey: Princeton University Press; 1994.
[36] Karanfil F. Energy consumption and economic growth revisited: does the size of unrecorded economy matter? Energy Policy 2008;36:3019–25.
[37] Dickey DA, Fuller WA. Likelihood ratio statistics for autoregressive processes. Econometrica 1981;49:1057–72.
[38] Phillips PCB, Perron P. Testing for unit root in time series regression. Biometrica 1988;75:335–46.
[39] Granger CWJ. Some recent developments in a concept of causality. Journal of Econometrics 1988;39:199–211.
[40] Lutkepohl H. Non-causality due to omitted variables. Journal of Econometrics 1982;19:367–78.
[41] Koop G, Pesaran MH, Potter SM. Impulse response analysis in nonlinear multivariate models. Journal of Econometrics 1996;74:119–47.
[42] Pesaran MH, Shin Y. Generalized impulse response analysis in linear multivariate models. Economics Letters 1998;58:17–29.